ADVANCES IN APPLIED MATHEMATICS 13, 454–461 (1992)

Characterizations of Generators for Modified de Bruijn Sequences

GREGORY L. MAYHEW

Hughes Aircraft Company, P.O. Box 2346, Fullerton, California 92633

AND

SOLOMON W. GOLOMB*

University of Southern California, Electrical Engineering 504 A, Los Angeles, California 90089

Order n modified de Bruijn sequences are created by removing a single zero from the longest run of zeros in period 2^n de Bruijn sequences. The M sequences are then the natural linear subset of modified de Bruijn sequences. Recursions which are the nonlinear duals to primitive polynomials over $GF(2)$ are developed. Data is presented for $4 \leq n \leq 6$. © 1992 Academic Press, Inc.

I. INTRODUCTION

Algebraically constructed binary sequences with randomness properties have applications in logic synthesis [1], coding theory [2], cryptography [3], and spread spectrum communications [4]. The *order n de Bruijn sequences* are the $2^{2^{n-1}-n}$ period 2^n binary sequences generated recursively using a n-stage feedback shift register [5]. Golomb provides a detailed treatment of the general properties of shift register sequences [6]. For $n \geq 4$, the de Bruijn sequences exhibit the balance, run, and span n randomness properties [7] and have linear spans greater than half the sequence length [8]. Removing a single zero from the longest run of zeros in a de Bruijn sequence produces the corresponding *modified de Bruijn sequence*. The M sequences are then the natural undisguised linear subset of modified de Bruijn sequences. The feedback functions which produce M sequences correspond to primitive (linear) polynomials over $GF(2)$. The (nonlinear) feedback function characteristics of the remaining order n modified de Bruijn sequences are unknown. This paper presents feedback function

*Dr. Golomb's research was supported in part by the United States National Security Agency under Grant No. MDA 904-91-H-0009.

0196-8858/92 $9.00

distributions for orders 4 through 6. Some theoretical characterizations of the nonlinear feedback functions are also provided.

II. Feedback Function Polynomial Distributions

The set of all order n de Bruijn sequences, $DS(n)$, are produced by an n stage feedback shift register. The next content of the least significant stage x_1 is computed as some feedback function $x_n \oplus g(x_{n-1} \cdots x_1)$ of the current values, where $\oplus$ denotes addition over GF(2). The function $g(x_{n-1} \cdots x_1)$ is easily represented by a truth table. The *weight* $w(g)$ of the feedback function is the number of logical ones (Hamming weight) among the 2^{n-1} entries in the truth table of the feedback function $g(x_{n-1} \cdots x_1)$. Fredricksen has characterized the weight classes for truth tables which produce $DS(n)$ [9].

Odd Weight Theorem. *There exists $S \in DS(n)$ with truth table weight w for every odd w between $Z(n) - 1$ and $2^{n-1} - Z^*(n) + 1$, inclusive.*

The bounds use the number of cycles from the pure cycling register $Z(n)$ and the number of cycles from the Complementing Cycling Register $Z^*(n)$,

$$Z(n) = \frac{1}{n} \sum_{\substack{d|n \\ \text{all } d}} \phi(d) 2^{n/d}, \qquad Z^*(n) = \frac{1}{2n} \sum_{\substack{d|n \\ \text{odd } d}} \phi(d) 2^{n/d},$$

where $\phi(n)$ is the Euler totient function [10].

Let $mDS(n)$ denote the set of all order n modified de Bruijn sequences.

Theorem 1. *There exists $S \in mDS(n)$ with truth table weight w for every even w between $Z(n) - 2$ and $2^{n-1} - Z^*(n)$, inclusive.*

Proof. The all zero state becomes its own successor by removing a zero from the longest run of zeros. Equivalently, the all zero state becomes its own successor by changing $g(0) = 1$ to $g(0) = 0$, thus decreasing the weight by one. □

The feedback functions for modified de Bruijn sequences are obtained by applying a logic reduction technique to each truth table. The truth tables were generated by computer using specialized tests to expedite the identification of degenerate truth tables that did not produce full period cycles [11]. The appropriate logic reduction technique is Reed Muller decoding which is based on Galois field arithmetic rather than Karnaugh maps which are based on Boolean algebra [12]. Obviously, each truth table producing a modified de Bruijn sequence corresponds to a unique feed-

TABLE I
Feedback Function Characteristic Polynomials
for Order 4 Modified de Bruijn Sequences

1. $x_4 \oplus x_3 \oplus 1$
2. $x_4 \oplus x_3 \oplus x_2 \oplus x_3x_2 \oplus x_3x_1 \oplus x_2x_1 \oplus 1$
3. $x_4 \oplus x_2 \oplus x_1 \oplus x_3x_2 \oplus x_3x_1 \oplus x_2x_1 \oplus 1$
4. $x_4 \oplus x_1 \oplus 1$
5. $x_4 \oplus x_3 \oplus x_3x_2 \oplus x_2x_1 \oplus 1$
6. $x_4 \oplus x_3 \oplus x_2 \oplus x_3x_1 \oplus 1$
7. $x_4 \oplus x_2 \oplus x_1 \oplus x_3x_1 \oplus 1$
8. $x_4 \oplus x_1 \oplus x_3x_2 \oplus x_2x_1 \oplus 1$
9. $x_4 \oplus x_2 \oplus x_3x_1 \oplus x_2x_1 \oplus 1$
10. $x_4 \oplus x_3 \oplus x_2 \oplus x_2x_1 \oplus 1$
11. $x_4 \oplus x_2 \oplus x_3x_2 \oplus x_3x_1 \oplus 1$
12. $x_4 \oplus x_2 \oplus x_1 \oplus x_3x_2 \oplus 1$
13. $x_4 \oplus x_3 \oplus x_2 \oplus x_1 \oplus x_3x_1 \oplus x_2x_1 \oplus 1$
14. $x_4 \oplus x_3 \oplus x_2 \oplus x_1 \oplus x_3x_2 \oplus x_3x_1 \oplus 1$
15. $x_4 \oplus x_3 \oplus x_2 \oplus x_3x_2 \oplus 1$
16. $x_4 \oplus x_2 \oplus x_1 \oplus x_2x_1 \oplus 1$

back function. Reed Muller decoding produces a unique function because the truth tables do not contain errors which must be corrected. An order $n - 1$ Reed Muller decoding is applied to $g(x_{n-1} \cdots x_1)$. In order $n - 1$ Reed Muller decoding, the implicants are x_{n-1} through x_1 and 1, which are linear, and all possible products of x_{n-1} through x_1, which are nonlinear. Complemented variables are not present in any implicant.

The feedback function polynomials for order 4 are presented in Table I. Note that the subscript notation is preferred so that cross products do not collapse into incorrect higher degree linear polynomial terms (i.e., $x_5x_4x_1$

TABLE II
Implicant Class Distribution for Order 4
Modified de Bruijn Sequences

Number of implicants	Number of sequences
3	2
5	10
7	4

TABLE III
Implicant Class Distribution for Order 5
Modified de Bruijn Sequences

Number of implicants	Number of sequences
3	2
5	96
7	492
9	884
11	494
13	76
15	4

looks like x^{10}). The number of terms in the feedback function polynomials for orders 4 through 6 are presented in Tables II through IV, respectively. Note the Gaussian nature of the distributions. Every feedback function polynomial includes the term $\oplus 1$. Much literature has been devoted to trinomials—the minimal shift register implementation of an M sequence [13].

TABLE IV
Implicant Class Distribution for Order 6
Modified de Bruijn Sequences

Number of implicants	Number of sequences
3	2
5	246
7	11,238
9	198,204
11	1,562,562
13	6,444,000
15	14,773,700
17	19,559,816
19	15,288,166
21	7,081,094
23	1,893,854
25	275,052
27	20,294
29	628
31	8

III. Theoretical Results

This section presents the theoretical results corresponding to the statistical results of the previous section. All sequences considered have terms in $GF(2)$. The cases of order n for $1 \leq n \leq 3$ are considered trivial because the only de Bruijn sequences for these orders are obtained directly from M sequences.

Let $\tau(n)$ denote the number of terms or implicants in the feedback function characteristic polynomial producing an order n modified de Bruijn sequence. Let $\delta(n)$ denote the degree of nonlinearity of this characteristic polynomial (or recursion). For example, the order 6 linear recursion $x_6 \oplus x_5 \oplus x_2 \oplus x_1 \oplus 1$ corresponding to the primitive pentanomial $x^6 \oplus x^5 \oplus x^2 \oplus x^1 \oplus 1$ has $\tau(6) = 5$ and $\delta(5) = 1$. Similarly, the order 6 nonlinear recursion $x_6 \oplus x_1 \oplus x_5x_4x_1 \oplus x_5x_3x_2x_1 \oplus 1$ has $\tau(6) = 5$ and $\delta(6) = 4$.

Lemma 2. $g(2^{n-1} - 1) = 1$.

Proof. If $x_n \oplus g(x_{n-1} \cdots x_1) = 1$ at the all ones state, then the all ones state is its own successor and not part of the modified de Bruijn cycle. Thus $g(01 \cdots 1) = 1$. □

Theorem 3. $\tau(n) = 1 \bmod 2$.

Proof. At the all ones state, each variable x_j in an implicant is evaluated at 1, so each implicant contributes $\oplus 1$ to the polynomial $x_n \oplus g(x_{n-1} \cdots x_1) \oplus 1$. Since $g(2^{n-1} - 1) = 1$, $g(2^{n-1} - 1)$ must be represented by an odd number of terms. □

For $n > 2$, if the feedback function does not make explicit use of all $n - 1$ variables, the corresponding shift register produces an even number of cycles [6].

Theorem 4. $\delta(n) \leq n - 2$.

Proof. The feedback function producing an order n modified de Bruijn sequence has two cycles, lengths $2^n - 1$ and 1, so the implicant $x_{n-1} \cdots x_2x_1$ is not in the feedback function. In the order $n - 1$ Reed Muller decoding, $x_{n-1} \cdots x_2x_1$ is the only implicant with $\delta(n) = n - 1$. □

Theorem 5. $\tau(n) \leq 2^{n-1} - 1$.

Proof. The $GF(2)$ logic reduction of an order n modified de Bruijn truth table has 2^{n-1} implicants available. The feedback polynomial never contains $x_{n-1} \cdots x_2x_1$ and the number of implicants is always odd. □

THEOREM 6. $3 \leq \tau(n)$.

Proof. The implicants x_n and 1 are always present linearly. The truth tables producing modified de Bruijn sequences have nonzero weight. The Reed Muller decoding of $g(x_{n-1} \cdots x_1)$ with nonzero weight produces at least one implicant. □

For sequence $S = \{s_0, s_1, \ldots, s_{k-1}\}$, the *reverse sequence* $rS = \{s_{k-1}, \ldots, s_1, s_0\}$. Two sequences S_1 and S_2 are *equivalent*, $S_1 = S_2$, if one is a cyclic shift of the other. For $n > 2$, $rS \neq S$ [14].

For M sequences, the characteristic polynomials of S and rS are related by the linear reciprocal transformation. When $f(x)$ is a polynomial of degree n over $GF(2)$, the reciprocal polynomial $f^*(x)$ is given by $f^*(x) = x^n f(1/x)$. This transformation maps variable x^{n-j} into variable x^j. The reciprocal transformation concept carries over directly from the linear shift register recursions to nonlinear shift register recursions. In developing the $GF(2)$ logic reduction, subscripts replaced superscripts so that nonlinear implicants are clearly distinguishable. The nonlinear reciprocal transformation maps variable x_{n-j} into variable x_j for every variable in an implicant, where the variables x_0 and 1 are equivalent. Note that the nonlinear reciprocal transformation preserves $\delta(n)$ and $\tau(n)$.

Let $\gamma(j, n)$ denote the number of order n modified de Bruijn sequences with $\tau(n) = j$.

THEOREM 7. *For* $n > 2$, $\gamma(j, n) = 0 \bmod 2$.

Proof. The reciprocal transformation creates a nonlinear polynomial with the time index reversed relative to the original polynomial and therefore produces the reverse sequence. The nonlinear reciprocal transformation preserves $\tau(n)$. For $n > 2$, $rS \neq S$. Hence S and rS exist as a distinct pair whose recursions have equal $\tau(n)$. □

Let $\lambda(j, n)$ denote the number of order n modified de Bruijn sequences with $\delta(n) = j$.

THEOREM 8. *For* $n > 2$, $\lambda(j, n) = 0 \bmod 2$.

Proof. The reciprocal transformation creates a nonlinear polynomial with the time index reversed relative to the original polynomial and therefore produces the reverse sequence. The nonlinear reciprocal transformation preserves $\delta(n)$. For $n > 2$, $rS \neq S$. Hence S and rS exist as a distinct pair whose recursions have equal $\delta(n)$. □

The reciprocal transformation process enables a second modified de Bruijn sequence to be created from any known generator function. This process is easily applied to the examples herein. The linear recursion $x_6 \oplus x_5 \oplus x_4 \oplus x_1 \oplus 1$ is the reciprocal of the linear recursion $x_6 \oplus x_5 \oplus x_2 \oplus x_1 \oplus 1$. The nonlinear recursion $x_6 \oplus x_5 \oplus x_5x_2x_1 \oplus x_5x_4x_3x_1 \oplus 1$

is the reciprocal of the nonlinear recursion $x_6 \oplus x_1 \oplus x_5x_4x_1 \oplus x_5x_3x_2x_1 \oplus 1$. Similarly, in Table I the following pairs are reciprocals: 1 and 4; 2 and 3; 5 and 8; 6 and 7; 9 and 11; 10 and 12; 13 and 14; 15 and 16. Thus the symmetry group $G_2 = \{e, r\}$ partitions the order n modified de Bruijn sequences into sets of two pairwise inequivalent sequences (e is the identity operator).

For large n, an M sequence characteristic polynomial can be obtained by known construction methods, by systematic searching, or by trial and error. In the latter case, a polynomial over $GF(2)$ is selected at random and then tested for primitivity. The chances of success are 1 in n trials. The chances of picking a nonlinear recursion that will produce an order n modified de Bruijn sequence are 1 in 2^{n-3} trials.

Let $LF(n)$ denote the number of valid candidates for $x_n \oplus g(x_{n-1} \cdots x_1) \oplus 1$.

THEOREM 9. $LF(n) = 2^{2^{n-1}-3}$.

Proof. The implicants x_n and 1 are always present and the implicant $x_{n-1} \cdots x_2x_1$ is never present. The function must be represented by an odd number of the remaining $2^{n-1} - 2$ implicants in an order $n - 1$ Reed Muller decoding. The number of odd subsets of a set with cardinality y is 2^{y-1}. □

THEOREM 10. $LF(n)/mDS(n) = 2^{n-3}$.

IV. CONCLUSIONS

By considering period $2^n - 1$ instead of period 2^n, the well-known M sequences become the linear subset of the modified de Bruijn sequences. The unique recursions provided by the $GF(2)$ logic reduction via Reed Muller decoding bring the nonlinear theory into alignment with the familiar linear theory. Specifically, the recursions have an odd number of terms between well-defined limits and reverse sequences are related by recursions with equal $\delta(n)$ and $\tau(n)$. The nonlinear recursions are the nonlinear duals to primitive polynomials over $GF(2)$.

REFERENCES

1. O. S. ROTHAUS, On "bent" functions, *J. Combin. Theory Ser. A* **20** (1976), 300–305.
2. F. J. MACWILLIAMS AND N. J. A. SLOANE, "The Theory of Error Correcting Codes," North Holland, New York, 1977.
3. W. MEIER AND O. STAFFELBACH, Nonlinearity criteria for cryptographic functions, *in* "Advances in Cryptology: Proceedings of EUROCRYPT '89," Springer-Verlag, New York/Berlin, 1989.

4. M. SIMON, J. OMURA, R. SCHOLTZ, AND B. LEVITT, "Spread Spectrum Communications," Vol. I, Computer Sci., Rockville, MD, 1985.
5. N. G. DE BRUIJN, A combinatorial problem, *in* "Kononklijke Nederlands Akademi van Wetenchappen, Proceedings," Vol. 49 (Part 2) (1946), 758–764.
6. S. W. GOLOMB, "Shift Register Sequences," 2nd ed., Aegean Park Press, Laguna Hills, CA, 1982.
7. S. W. GOLOMB, On the classification of balanced binary sequences of period $2^n - 1$, *IEEE Trans. Inform. Theory* **IT-26**, No. 6, (1980), 730–732.
8. A. H. CHAN, R. A. GAMES, AND E. L. KEY, On the complexities of de Bruijn sequences, *J. Combin. Theory Ser. A* **33** (1982), 233–246.
9. H. FREDRICKSEN, A survey of full length nonlinear shift register cycle algorithms, *SIAM Rev.* **24** (1982), 195–229.
10. I. NIVEN AND H. S. ZUCKERMAN, "The Theory of Numbers," Wiley, New York, 1980.
11. G. L. MAYHEW, "Statistical Properties of Modified de Bruijn Sequences," Ph.D. dissertation, University of Southern California, December 1987.
12. I. S. REED, A class of multiple error correcting codes and the decoding scheme, *IRE Trans. Inform. Theory* **4** (1954), 38–49.
13. N. ZIERLER AND J. BRILLHART, On primitive trinomials (mod 2), *Inform. and Control* **13** (1968), 541–554.
14. T. ETZION AND A. LEMPEL, On the distribution of de Bruijn sequences of given complexity, *IEEE Trans. Inform. Theory* **IT-30**, No. 4, (1984), 611–614.